ABSTRACT

Risk assessment and diagnosis of a complex disorder often requires measuring an underlying quantitative phenotype. Association studies in unrelated populations can implicate genetic factors contributing to disease risk, and experiments using pooled DNA provide a less costly but necessarily less powerful alternative to methods based on individual genotyping. Although the sample sizes required for pooling and individual genotyping studies have been compared in certain instances, general results have not been reported in the context of association studies, nor have there been clear comparisons of pooling based on quantitative and qualitative (affected/unaffected) phenotypes. Here we use exact numerical calculations and analytical approximations to examine the sample size requirements of association tests for quantitative traits and affected-unaffected studies using pooled DNA. We show, in analogy with selection experiments, that the optimal design for virtually any quantitative phenotype is to pool the top and bottom 27% of individuals, regardless of marker frequency or inheritance mode; this design requires a population only 24% larger than that required for individual genotyping. Furthermore, this design is approximately four times more efficient than typical affected-unaffected studies of DNA pooled from individuals classified as affected or unaffected.